Perception, 1998, volume 27, pages 257-272

# Stereoscopic discrimination of interval and ordinal depth relations on smooth surfaces and in empty space

J Farley Norman
Department of Psychology, Western Kentucky University, Bowling Green, KY 42101, USA;
e-mail: Farley.Norman@wku.edu
James T Todd
Department of Psychology, Ohio State University, Columbus, OH 43210, USA
Received 21 October 1996, in revised form 7 July 1997

**Abstract.** In a series of three experiments, observers judged the perceived relative depths of small probe dots, which could be presented in empty space or attached to a smoothly curved surface. Discriminations of ordinal depth were found to be more precise than discriminations of depth intervals. The amount of separation in the projected image between the locations in depth was also manipulated. Performance was higher when observers evaluated the depth relationships between nearby points in the projected images, and lower when the points were more widely separated. This effect was most pronounced when there was a continuous surface in between the points, suggesting that accurate knowledge of the three-dimensional structure of surfaces is primarily limited to relatively small local neighborhoods.

## 1 Introduction

Ever since Wheatstone's (1838) invention of the stereoscope, many researchers have studied the perception of stereoscopic depth. Many of these investigations have examined the *precision* of human observers' knowledge of depth relationships. A common paradigm has been to show two or three isolated elements in the visual field (bars, lines, spots of light), and ask observers to adjust them to equidistance (for example, see Ogle 1956; Enright 1991) or to discriminate whether one element is either closer or farther than another (Blakemore 1970; Westheimer 1979). It became apparent early on that observers can perceive near/far depth relationships with remarkable precision in some circumstances. For example, von Helmholtz (1867/1962) viewed three vertical pins at a distance of 340 mm, and adjusted the center pin in depth, until the three pins appeared to lie in the same depth plane. According to Helmholtz, he never made an error larger than half the width of one of those pins. Westheimer and McKee (1978, 1979, 1980) later showed that observers exhibit stereoacuities similar to Vernier hyperacuities, and can reliably detect depth differences smaller than the width of a single retinal photoreceptor.

However, these extensive studies over the past 150 years have also indicated that the stereoscopic perception of depth differences is not always precise, and depends greatly upon how far away two targets are placed in depth relative to the plane of fixation (Westheimer and McKee 1978; McKee et al 1990), and on the angular separation of two targets in the visual field (Wright 1951; Ogle 1956; Enright 1991). For widely spaced objects, better performance is obtained with successive alternating eye fixation on each target than with sustained fixation on either target alone (Ogle 1956; Enright 1991, 1996). This result suggests that the ability to detect small differences in depth is limited to mechanisms that have receptive fields in the central parts of the visual field.

Traditionally, the slight differences between the two eye's views of the world have been quantified in terms of relative binocular disparities, where disparity is defined as the difference in convergence angles that would be needed to fixate each of two targets at differing locations in depth (Ogle 1950; Foley 1978; Cormack and Fox 1985). A schematic illustration of a binocular viewing situation is shown in figure 1.

Alternatively, one can examine the differences that arise in the two-dimensional (2-D) projections of the three-dimensional (3-D) scene that are presented to the eyes (figure 2; also see Cormack and Fox 1985). In general, knowledge of either the relative binocular disparity or of the resulting image differences alone is insufficient to recover the precise magnitudes of depth intervals in the environment. A complete solution to the problem is possible only if the visual system can accurately measure the magnitude of the disparity, and either estimate the viewing distance to one of the targets by some means (eg the convergence state of the eyes) or utilize the vertical disparities that occur in the retinal images when eccentric targets are present within the scene (Longuet-Higgins 1982; also see Foley 1978).

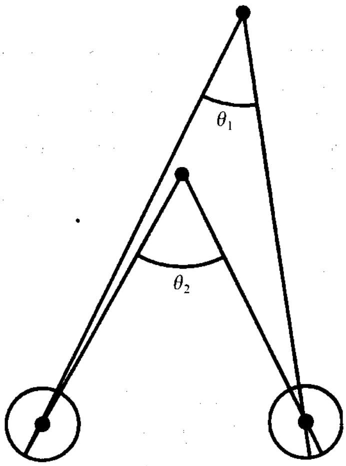

**Figure 1.** The geometry of stereopsis. Binocular viewing of two targets differing in their locations in depth. The relative binocular disparity is traditionally defined as the difference between the two angles $\eta = \theta_2 - \theta_1$.

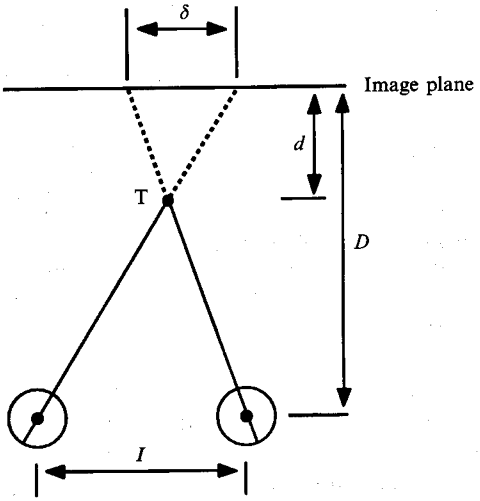

**Figure 2.** Binocular viewing of a target T in front of the image plane. The dashed lines indicate the projected positions of the target in the left and right half-images of the stereogram. The image disparity $\delta$ indicates the difference in positions across the two half-images. $\delta$ is zero when T lies in the image plane and increases in magnitude as the depth interval $d$ between T and the image plane becomes larger. The amount of disparity also depends upon the interpupillary distance between the observers' eyes ($I$) and the distance to the image plane $D$.

In contrast to the computational difficulty of accurately measuring the size of depth intervals, determining the orderings of objects in depth is relatively straightforward. Figure 3 shows the image differences that arise from viewing two objects at differing depths. In this case, an observer needs to only determine the sign of the difference between the two image disparities—ie the target with the larger image disparity is closest to the observer (the situation is reversed for targets located behind the projection plane). Based upon his analysis of what information could potentially be provided by stereoscopic disparity, Ogle (1958, page 756) wrote that "the stereoscopic depth, as experienced from the images of a number of different objects at different distances in space, alone may indeed provide only a 'rank-order' scheme of depth". Similar conclusions about the possibility of human observers' knowledge not being entirely accurate with respect to the perceived magnitudes of depth intervals have been expressed by Todd and Reichel (1989) for surfaces defined by image shading, and by Todd and Norman (1991) and Norman and Todd (1993) for the perception of 3-D structure from motion.

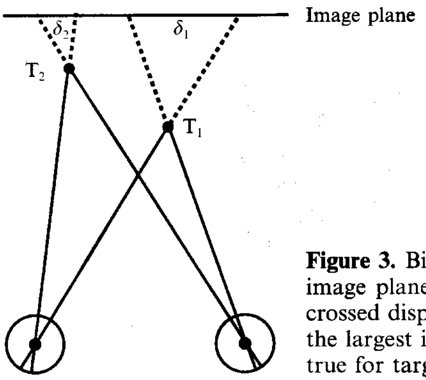

**Figure 3.** Binocular viewing of two targets $T_1$ and $T_2$ in front of the image plane. $\delta_1$ and $\delta_2$ indicate the resulting image disparities. For crossed disparities located in front of the image plane, the point with the largest image disparity is closest to the observer. The converse is true for targets located behind the image plane.

The currently available psychophysical evidence suggests that human observers' ability to perceive the magnitude of stereoscopically defined depth intervals may be rather poor, in contrast to the remarkable performance often obtained for judgments of depth order. For brief stimulus presentations, McKee et al (1990) found that the Weber fraction for disparity ranged from 10%–20%, while the Weber fractions for discriminations of monocular width judgments in the frontoparallel plane were only 2%–3%. Analogous findings for the perception of 3-D length were reported by Norman et al (1996). Finally, for the perception of depth intervals between separated regions on smooth surfaces, Norman and Todd (1996) found Weber fractions that ranged from 12% to 40%, depending on such factors as the magnitude of the standard interval and the amount of separation between the two regions on the surfaces of the objects.

The purpose of the current set of experiments was twofold. First, we sought to directly compare the perception of ordinal and interval depth, for the same observers and similar stimuli. Many experiments have been concerned with the perception of ordinal depth (apparent equidistance, near/far discriminations, etc) or the perception of the magnitudes of depth intervals, but these two abilities have rarely, if ever, been examined together. In addition, many previous studies of the perception of ordinal stereoscopic depth have involved the manipulation of two or three isolated elements in the visual field, often viewed in a dark room (for an interesting exception where planar, textured surfaces have been used, see Enright 1996). We also wanted to examine the perception of ordinal depth relationships on continuous, smoothly curved objects that were perceptually well defined by many sources of optical information, including stereoscopic disparity, texture gradients, and image shading.

## 2 Experiment 1

Suppose that an observer views two small dots at differing locations, one at a standard location in depth while the other is either closer or farther away. If the observer is asked to judge which one appears closer in depth, he can fixate on one of the dots, and determine a response from the sign of the image disparity difference (see figure 3). One can vary the amount of depth separation over time and calculate the minimum depth difference that is reliably detectable (for an example of this paradigm, see Westheimer and McKee 1980). Now suppose, however, that the dot marking the standard location in depth is removed and that all other aspects of the display are identical. In this case, the observer views a single point hovering in depth, and is asked to judge whether it is nearer or farther than a standard position that is no longer explicitly visible in the stimulus. Is this task possible, on the basis of stereoscopic vision? If there is another object in the scene, the observer can perform this task by evaluating the magnitude of the resulting depth *interval* (between the test point and the reference object) as it changes over time (for an example of this paradigm, see McKee et al 1990).

The purpose of this first experiment was to compare performance on these two types of tasks, which in one sense are quite similar, but which also place different computational demands upon the observer. The detection of "nearer" vs "farther" when two targets are simultaneously present is computationally much simpler (see figure 3) than that of evaluating the changing depth intervals or disparities that take place when only a single target is visible. In this case, an accurate estimate of the magnitude of the disparity is needed to make a correct response. A secondary purpose of the experiment was to verify the effect of separation in the visual field on the observers' ordinal depth judgments when the targets were viewed in a well lit environment. Previous investigators (Wright 1951; Ogle 1956; Enright 1991) have found large effects of increasing separation for simultaneously presented targets, but their experiments were conducted in dimly lit or darkened rooms.

### 2.1 *Method*

2.1.1 *Apparatus.* The optical patterns were created and displayed on a Silicon Graphics Crimson VGXT workstation. The stereoscopic half-images were presented with the use of LCD (liquid crystal) shuttered glasses that were synchronized with the monitor's refresh rate. The left and right views of a stereo pair were displayed at the same position on the monitor screen, but they were temporally offset. The left and right lenses of the LCD glasses shuttered synchronously with the display so that each view of the stereo pair was seen only by the appropriate eye. The CRT was refreshed at 120 Hz. Thus each view of a stereoscopic half-image was updated at a rate of 60 Hz. The viewing distance was 76.0 cm, such that the 1280 pixel wide by 1024 pixel high display screen subtended 25.2 by 20.3 deg of visual angle.

2.1.2 *Stimulus displays.* Depending on the condition, one or two white circles (0.5 cm, or 22.6 min arc diameter) were presented stereoscopically in front of a yellow grid. This grid subtended 15 deg both horizontally and vertically, and was composed of 15 evenly spaced vertical bars and 15 evenly spaced horizontal bars. The background grid was placed at the same depth as the computer screen. The spatial position of each circle was randomly varied across trials. An example stereogram is shown in figure 4. All of the displays were generated with the appropriate perspective for a 76 cm viewing distance. Each eye's view was computed on the basis of the observers' individual interpupillary distances. Anti-aliasing hardware was used to render the stimulus displays, so that the spatial locations of the centers and edges of the circular spots were more precisely defined than pixel resolution—the effective resolution was at least 0.1 pixel. The stereograms were viewed in full-cue conditions in a lighted room. The observers could see the computer monitor and its placement in depth. Information about the viewing distance to the monitor was thus available from the convergence state of the observers' eyes.

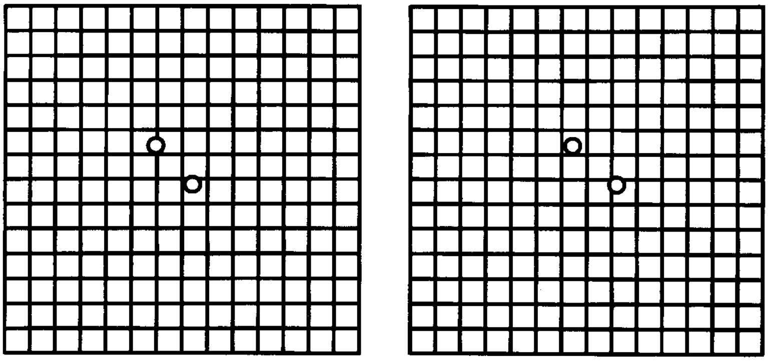

**Figure 4.** A stereogram similar to those used in experiment 1, depicting two circular spots hovering in depth. The right spot is closer than the left. This stereogram is designed for crossed free fusion.

2.1.3 *Procedure.* The observers' task on each trial was to judge whether a test spot was nearer or farther than a standard position in depth, which was located 3.0 cm in front of the plane of the computer screen (approximately a disparity of 11.4 min arc; this was different for each observer owing to their individual interpupillary distances). Each correct response was followed by a short auditory beep. There were three different experimental conditions. In the *ordinal-near* condition, the standard depth location was marked by a visible spot, and a second test spot was placed either in front of or behind this standard depth. The spots were randomly placed horizontally and vertically, but were constrained so that the separation of their centers in the projected image was greater than 23 min arc (to prevent overlap), but less than 3 deg. The average separation over trials was approximately 1.7 deg. The observer indicated which spot appeared closer in depth—the right spot in the image or the left spot—by pressing one of two buttons on the workstation mouse. It is important to keep in mind that the observers did not know which of the two spots—the right or the left—was located at the standard depth. Accurate performance for the task required attending to both spots simultaneously. The *ordinal-far* condition was identical, except that the spot separation in the projected image was always greater than 3 deg, but less than 11.3 deg. The average separation over trials in this condition was about 7.1 deg. In the *interval* condition, the spot marking the standard depth location was removed—the displays on each trial were otherwise identical to those of the ordinal-near and ordinal-far conditions. The observer pressed one button if the test spot appeared closer than the now implicit standard depth location and pressed a different button if the test spot appeared farther than the implicit standard. This was effectively equivalent to judging whether the depicted depth interval between the single spot and the background grid was greater or less than 3.0 cm. The observers were given 20 practice trials at the beginning of each block. Since feedback was provided after every trial, they could learn the standard location in depth prior to the start of the experimental trials. In all conditions, the observers could view each stimulus presentation for as long as they desired, until they were ready to make a response.

For each of the three conditions, there were eight possible test positions in depth: four closer to the observer than the standard and four farther than the standard. In particular, the test was either closer or farther than the standard by 0.04, 0.12, 0.20, or 0.28 cm (based on the average interpupillary distance for our observers of 6.17 cm, these depth differences correspond roughly to disparities of 9.5, 28.7, 47.8, and 67.0 s arc, respectively). In each experimental block, there were 20 practice trials plus 50 trials for each of the eight test locations in depth. Therefore, each block consisted of 420 trials. The observers participated in two blocks for each of the three conditions. At the end of the experiment, a total of 100 trials had been obtained for each combination of the eight test depth locations and three experimental conditions.

2.1.4 *Observers.* The observers included both authors (JFN and JTT), and one other experienced psychophysical observer (HFN).

### 2.2 *Results*

Psychometric functions were estimated for each observer and condition by finding the cumulative normal distribution that best fit the data for each of the eight test locations in depth. This was done with the use of a probit analysis program developed by Foster and Bischof (1991). The depth difference thresholds for judging whether the test depths were closer or farther than the standard location were obtained by estimating the 75% and 25% points on the observers' psychometric functions, respectively. The standard deviations of the threshold estimates were calculated by a bootstrap procedure (Foster and Bischof 1991, with 1000 bootstrap iterations). Figure 5 shows, for each observer and condition, the average of the thresholds for judging "greater than" or "less than", expressed in terms

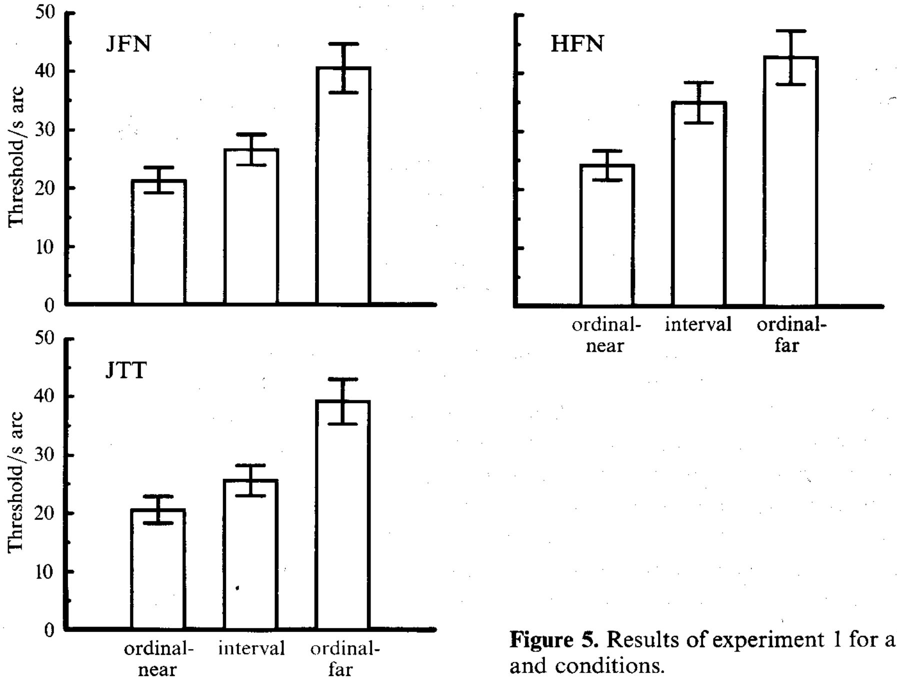

**Figure 5.** Results of experiment 1 for all observers and conditions.

of disparities. On average, the observers required a disparity of 22 s arc to discriminate ordinal depth relations when the two spots were close together. In the interval condition, the observers needed a disparity increment or decrement of 30 s arc relative to the implicit standard location (or mean test location over time) to detect that the depth interval between the test spot and the background was either greater or smaller in magnitude. Finally, when the observers were asked to discriminate ordinal relationships in depth between widely separated spots, the thresholds increased to over 40 s arc. Each half-height of the error bars indicates the size of the average standard deviation of the two threshold estimates for each condition. One can also express these thresholds as Weber fractions—what depth difference does one need to add or subtract to the standard depth in order to detect either the ordinal depth relationships or whether a depth interval has increased or decreased (cf McKee et al 1990). Table 1 shows Weber fractions for the conditions in the current experiment.

**Table 1.** Results of experiment 1. Weber fractions expressed as a percentage of the standard.

| | Ordinal-near | Interval | Ordinal-far |
|---|---|---|---|
| JFN | 3.2 | 3.9 | 6.0 |
| HFN | 3.3 | 4.8 | 5.8 |
| JTT | 2.8 | 3.4 | 5.3 |

Since the same test depth differences were used for both the ordinal and interval conditions, the frequencies of responding "greater than" were subjected to a $\chi^2$ analysis to determine if the observed differences were statistically significant. There were significant differences between the ordinal judgments (near) and the interval judgments for all observers ($\chi_7^2 = 16.1$ for observer JFN, 27.2 for HFN, and 29.6 for JTT, all $p < 0.05$). There were also large differences between the ordinal judgments at the near and far image separations ($\chi_7^2 = 36.5$ for observer JFN, 63.4 for HFN, and 148.4 for JTT, all $p < 0.05$).

Our results for the ordinal depth discrimination tasks confirm the earlier findings of Wright (1951), Ogle (1950, 1956), and Enright (1991) that detection of near/far relationships deteriorates considerably as the separation between two targets is increased. This phenomenon seems to be very robust, and occurs not only in darkened, impoverished visual environments, but also when the stimulus display is viewed in a well-lit room with many sources of information about viewing distance, placement of the computer monitor, etc. Indeed, the thresholds obtained in the current experiment for separations of 1.7 and 7.1 deg are not only qualitatively similar to those of Ogle (1956, figure 2) for alternating fixation, they are quantitatively similar as well. The thresholds obtained for the interval task are similar to those of McKee et al (1990, see figure 5) and Westheimer (1979, see figure 3). Our mean test location in depth over time was at a disparity of 11.4 min arc with respect to the background plane. For that value of standing disparity, the disparity increment threshold for observer SPM in McKee et al's study was about 2 min arc for a display duration of 150 ms, which dropped to 1 min arc for a longer duration of 1000 ms. Our observers' thresholds for this condition were about 30 s arc—presumably, the improvement in our experiment was due to the fact that our observers could view each display for as long as they wished, until they were ready to make a response. Our results also agree with the findings of Westheimer (1979), who obtained disparity increment thresholds of about 20 s arc for a standing disparity of 2 min arc.

All of our observers showed a consistent elevation of thresholds in the interval condition relative to those in the ordinal-near condition (see figure 5 and table 1). There was an increase of 24% for observers JFN and JTT, and an increase of 45% for observer HFN. If anything, these results are a conservative underestimate. In the interval condition, there was always a small separation in the visual field between the target spot and some of the lines and intersections comprising the grid. In contrast, the average angular separation between the spots in the ordinal-near condition was 1.7 deg. In that condition, we obtained thresholds of about 22 s arc, which agrees well with the results of Ogle (1956). However, the data of Westheimer and McKee (1980) suggest that if the two targets in the ordinal-near condition had been placed even closer together, the thresholds might have been as low as 10 s arc.

There are two distinct possibilities why performance might be lower for interval judgments. First, the reduced performance might be a reflection of the increased computational difficulty involved in calculating depth intervals. Judgments in the interval condition required observers to estimate the magnitude of the changing disparity between the test spot and the background grid. This was not the case for the ordinal judgments when the two targets were simultaneously present: the precise knowledge of magnitudes of disparity was not needed for accurate performance (see figure 3). The second possibility is that the poorer performance in the interval condition reflects the use of an implicit standard. Perhaps the presence of a visible standard increased the performance in the ordinal-near condition, while the lack of a visible standard in the interval condition decreased performance. It is possible that our observers were unable to form a reasonably precise mental representation of the standard depth interval. Experiment 2 was designed to test this hypothesis.

## 3 Experiment 2

### 3.1 *Method*

3.1.1 *Stimulus displays.* The stimulus patterns were similar to those of experiment 1. The stereograms depicted one or two depth intervals. Each interval was defined by the separation in depth between a pair of white circles, which appeared to the observers as hovering in depth in front of the grid placed in the plane of the computer screen. All other details were identical to those of experiment 1. An example stereogram is shown in figure 6.

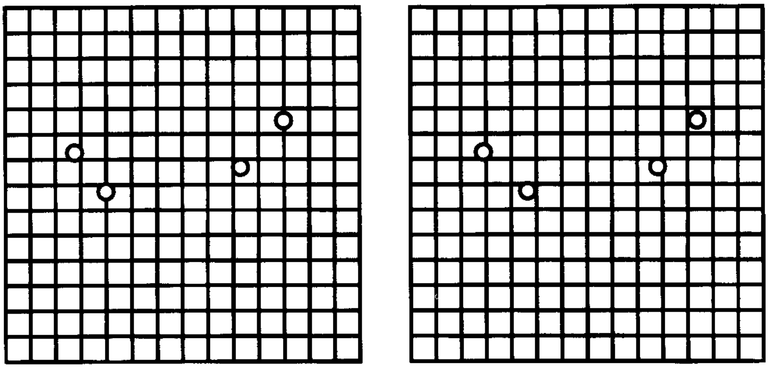

**Figure 6.** A stereogram similar to those used in experiment 2, depicting two depth intervals. The right interval is smaller in magnitude than the left. This stereogram is designed for crossed free fusion.

3.1.2 *Procedure.* The observers' task on each trial was to decide whether a test depth interval was greater or smaller in extent than a standard depth interval. A correct response was followed by a short auditory beep. In the *explicit* condition, the observer saw the standard and test depth intervals simultaneously during each trial. The observers' task was to indicate which of the intervals was greater in magnitude, the one on the left, or the one on the right. The observers were unaware which interval had the standard magnitude. For both standard and test, random offsets between 0.0 and 3.5 cm (approximately between 0 and 13.5 min arc disparity, given an average interpupillary distance of 6.17 cm) were applied on each trial to displace the intervals away from the background grid and towards the observer in depth. The observers made their responses by pressing different buttons on the workstation mouse. Presenting the standard and test intervals at different distances from the background was a necessary precaution to ensure that the observers compared the stimuli in terms of their relative depth interval magnitudes. If both intervals had been anchored at the same minimum depth then the observers could have discriminated between the two intervals using a strategy involving ordinal depth—ie the interval with the spot closest to the observer has the greatest interval magnitude. The *implicit* condition was identical to the explicit condition, except that the standard interval was not presented—only the test depth interval was visible. The observers formed a mental representation of the magnitude of the standard based upon the pattern of auditory feedback. A series of 20 practice trials were presented at the beginning of each block so that the observers had a clear impression of the standard depth interval prior to the start of the experimental trials. In this condition, the observers' task was to press one mouse button if the test depth interval appeared larger than the implicitly defined standard, and to press a different mouse button if the test interval appeared smaller than the standard. In all conditions, the observers could view each stimulus presentation for as long as they desired, until they were ready to make a response.

The magnitude of the standard depth interval was always 3.0 cm; the disparities of each circular spot defining the interval depended on the magnitude of the random offset on every trial. There were eight test depth intervals for each condition: four larger than the standard, and four smaller than the standard. In particular, for the explicit condition, the test intervals were greater and smaller than the standard by 4%, 12%, 20%, and 28%. For the implicit condition, the test intervals were greater and smaller than the standard by 2.5%, 7.5%, 12.5%, and 17.5%. In each experimental block, there were 20 practice trials plus 50 trials for each of the eight test intervals. Therefore, each block consisted of 420 trials. The observers participated in two blocks for each condition. At the end of the experiment, 100 trials had been obtained for each combination of the eight test depth interval magnitudes and two experimental conditions.

3.1.3 *Observers.* The displays were judged by the same three observers who had taken part in experiment 1.

### 3.2 *Results*

The observers' thresholds expressed as Weber fractions for both conditions are shown in table 2. Note in the table that, for each individual observer, the highest performance occurred in the implicit standard condition. The average Weber fraction for those judgments was 6.2%, whereas it jumped to over 10% when the standard was explicitly presented on each trial. Thus it would appear that the difference between interval and ordinal judgments in experiment 1 cannot be due to the absence of an explicit standard in the interval condition. Indeed, on the basis of the results of the present study, it is likely that the use of an explicit standard would have made the difference between interval and ordinal judgments even larger. There is other evidence to suggest, however, that this finding cannot be generalized to other types of psychophysical judgments. For example, in the discrimination of lifted weights, Wever and Zener (1928) and Fernberger (1931) found that the thresholds obtained with implicit standards were consistently higher than would otherwise be possible when an explicit standard is presented on every trial.

**Table 2.** Results of experiment 2. Weber fractions expressed as a percentage of the standard.

| | Explicit standard | Implicit standard |
|---|---|---|
| JFN | 11.7 ± 1.2 | 5.8 ± 0.6 |
| HFN | 10.4 ± 1.0 | 7.2 ± 0.7 |
| JTT | 10.2 ± 1.0 | 5.7 ± 0.6 |

Note: The precision for each threshold is printed to the right, and is the standard deviation of the threshold estimate as calculated by the bootstrap procedure (with 1000 bootstrap iterations) described by Foster and Bischof (1991).

Why should the effect of an explicit standard vary across different tasks? In order to better appreciate this issue, it is useful to consider the specific requirements for discriminating depth intervals in the present experiment. It is important to keep in mind that in the explicit standard condition observers did not know which of the two dot pairs on any given trial constituted the standard, so it was necessary for them to judge each pair separately. We will assume that there is some variance $\delta_s$ in the perceived depth interval between a pair of dots when it is examined over multiple occasions, and that judgments for different dot pairs are independent of one another. Thus, the variance $\delta_c$ in estimating the relative magnitude of two depth intervals would be governed by Thurstone's (1927) law of comparative judgment, Case V, such that: $\delta_c = \delta_s\sqrt{2}$. For the implicit standard condition, in contrast, the depth interval between a single pair of visible dots must be compared with a mental representation of the standard that is stored in memory. Suppose that this representation has a variance $\delta_m$ and that its value on each trial is independent of the visible target to which it must be compared. Under these conditions, the variance $\delta_c$ of the difference distribution would be governed by Thurstone's Case III, such that: $\delta_c = \sqrt{\delta_s^2 + \delta_m^2}$.

It is important to note from this analysis that the relative precision of observers' judgments for the two types of standard would depend on the relative magnitudes of $\delta_m$ and $\delta_s$. If $\delta_m > \delta_s$, then an implicit standard would produce a higher variance in observers' comparison judgments than would an explicit standard. Alternatively, if $\delta_m < \delta_s$, then the variance would be highest in the explicit condition. From the results of the present experiment, it seems reasonable to conclude that the second of these alternatives is applicable for the discrimination of depth intervals from binocular disparity. This suggests that the pattern of response feedback over a sequence of many trials allows observers to develop a mental representation of the standard that is defined more precisely (ie has a smaller variance) than the perceived depth interval between visible targets.

## 4 Experiment 3

In a previous series of experiments, Norman and Todd (1996) showed that the ability of observers to discriminate depth intervals between pairs of probe points is significantly impaired if the probe points are positioned on smoothly curved surfaces rather than empty space. One possible explanation of this finding is that the perceptual representation of smooth surfaces is primarily based on higher-order differential properties such as orientation or curvature (see also Rogers and Cagenello 1989), and that the computation of depth intervals between pairs of probe points in this type of data structure requires a process of integration over intervening surface regions. Suppose, however, that observers are asked to make ordinal judgments, for which the computation of depth intervals is no longer necessary. Would there still be an effect of intervening surface structure? Experiment 3 was designed to examine this issue.

### 4.1 *Method*

4.1.1 *Stimulus displays.* A set of 100 objects was generated by applying a series of sinusoidal distortions to a sphere at randomly selected positions and orientations (see Norman et al 1995; Todd and Norman 1995; Norman and Todd 1996). The surface of each object was defined by 1280 triangles arranged in a polygonal mesh. The transformation was accomplished by modulating this mesh in depth with a unidimensional sine wave. The surface was then rotated a random amount about all three Cartesian coordinate axes. The surface in its new orientation was again modulated sinusoidally, rotated randomly, etc, and this process was repeated ten times. The scaling factor of the modulation controls the amplitude of the bumps, dimples, and ridges of the resulting surface, while the number of iterations controls its apparent 'complexity'. The resulting objects were smoothly curved with no discontinuities, and by keeping track of each successive sinusoidal perturbation, we were able to obtain an analytically defined surface normal at each point. Two representative objects that were generated with this procedure are shown in figure 7.

The objects were presented with shading, texture, and stereoscopic disparities simultaneously. The surface shading was simulated with the use of a standard computer graphics reflectance model (see Todd and Mingolla 1983), in which the shading is partitioned into three components: an ambient component that is constant for all surface orientations, a diffuse (Lambertian) component that varies with the cosine of the angle between the surface normal and the direction of illumination, and a specular component that varies as a function of the surface normal, the direction of illumination, and the direction of view. For the displays in the present experiment, the specular component of the shading model was set to zero, simulating the reflectance properties of a matte surface. The Lambertian component of the shading model was 70%, while the ambient component was 30%.

The shading was consistent with a single simulated point-light source oriented with a slant of 28° and a 45° tilt up and to the left of the observers' line of sight. Texturing was applied to the surface by using a 2-D blue-and-white checkerboard pattern. Each polygon was first rotated to a frontoparallel orientation and then mapped onto a random region of the 2-D texture pattern, which ensured that equal areas of the surface contained equal amounts of texture (see Todd and Mingolla 1984).

Two local regions were highlighted on each surface by small red spheres (0.25 cm, or 11.3 min arc diameter). To ensure that the region under each spot was clearly visible and not too close to an object's occlusion boundary, we chose regions that had slants

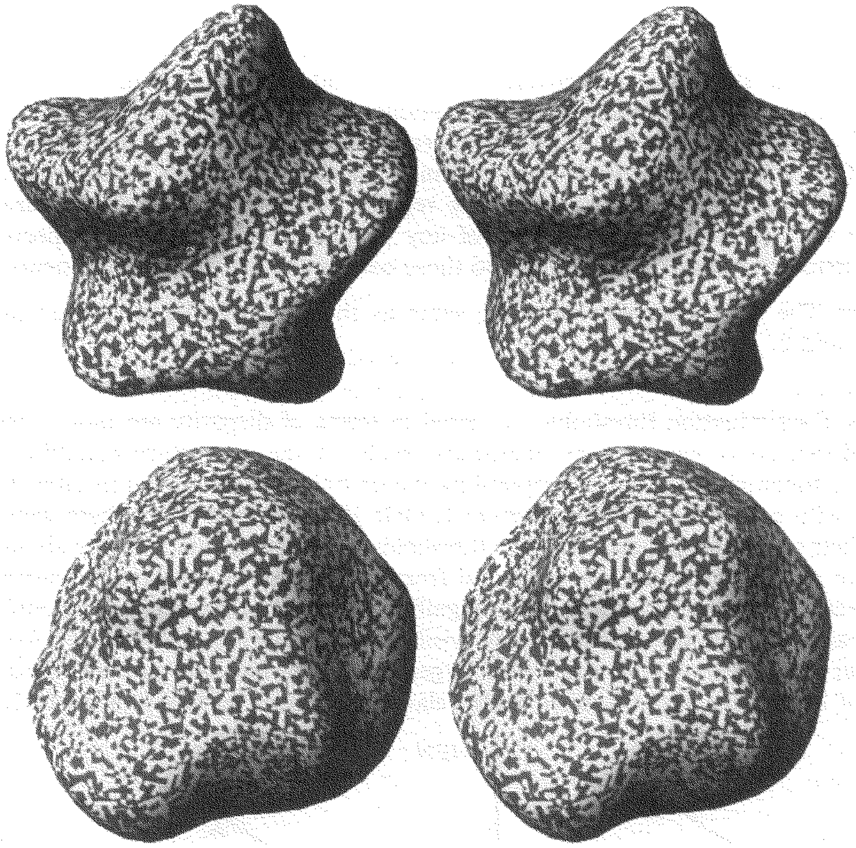

**Figure 7.** Stereograms of two of the randomly shaped objects used in experiment 3. These stereograms are designed for crossed free fusion.

less than or equal to 70°. There were two basic types of stimulus displays used in the current experiment. For the *near* condition, the two local regions were constrained so that their spatial separation was always 3.0 deg or less in the projected image. For the *far* condition, in contrast, the two regions were constrained so that their image separation would always be greater than 3.0 deg, but less than 11.3 deg. All other aspects of the stimulus generation were identical to those used in experiments 1 and 2.

4.1.2 *Procedure.* The experimental procedure was identical to that used in the ordinal conditions in experiment 1. On each trial an observer viewed one of the 100 possible 3-D objects in a random 3-D orientation. The observers were instructed to evaluate the relative depth order of the two highlighted local regions, and to indicate which of the regions was closer, the right or the left, by pressing one of two buttons on the workstation mouse. One of the two regions was located at a standard depth of 3.0 cm from the plane of the computer monitor (approximately a disparity of 11.4 min arc; this was different for each observer owing to their individual interpupillary distances), while the remaining test region was either closer or farther. In the near separation condition, the test regions were closer or farther than the standard depth location by 0.07, 0.21, 0.35, or 0.49 cm (based on the average interpupillary distance for our observers of 6.17 cm, these depth differences correspond roughly to disparities of 0.3, 0.8, 1.4, and 1.9 min arc for test spots farther than the standard and 0.3, 0.8, 1.4, and 2.0 min arc for test spots closer than the standard, respectively). In the far separation condition, the test regions were closer or farther than the standard depth by 0.15, 0.45, 0.75, or

1.05 cm (these depth differences correspond to disparities of 0.6, 1.8, 2.9, and 4.1 min arc for test spots farther than the standard and 0.6, 1.8, 3.0, and 4.2 min arc for test spots closer than the standard, respectively). There were 50 trials at each of the eight test depth locations for each condition plus an additional 20 practice trials, resulting in a total of 420 trials per block. Each observer participated in two experimental blocks. Therefore, a total of 100 trials were obtained for each test depth location and condition for all observers. All pairs of surface regions within a block of trials were unique—there were no repeat presentations of any particular pair. All other aspects of the experimental procedure were identical to those used in the preceding experiments.

4.1.3 *Observers.* The three observers were the same as those who had participated in experiments 1 and 2.

### 4.2 *Results*

The observers' discrimination thresholds expressed in terms of disparity are plotted in figure 8 (filled circles), along with the analogous results of the no surface conditions in experiment 1 (open circles). The corresponding Weber fractions are shown in table 3. In agreement with earlier findings (experiment 1; Ogle 1956; Enright 1991), there were significant effects of separation, such that the thresholds increased dramatically by about a factor of two as the separation was increased from near to far. A repeated-measures analysis of variance confirmed that there was a significant interaction between the amount of image separation and the presence of a visible surface ($F_{1,2} = 142.7$, $p < 0.01$), suggesting that the effects of separation were larger when a surface was present. Thus, relatively good knowledge about ordinal depth relationships on a surface is apparently

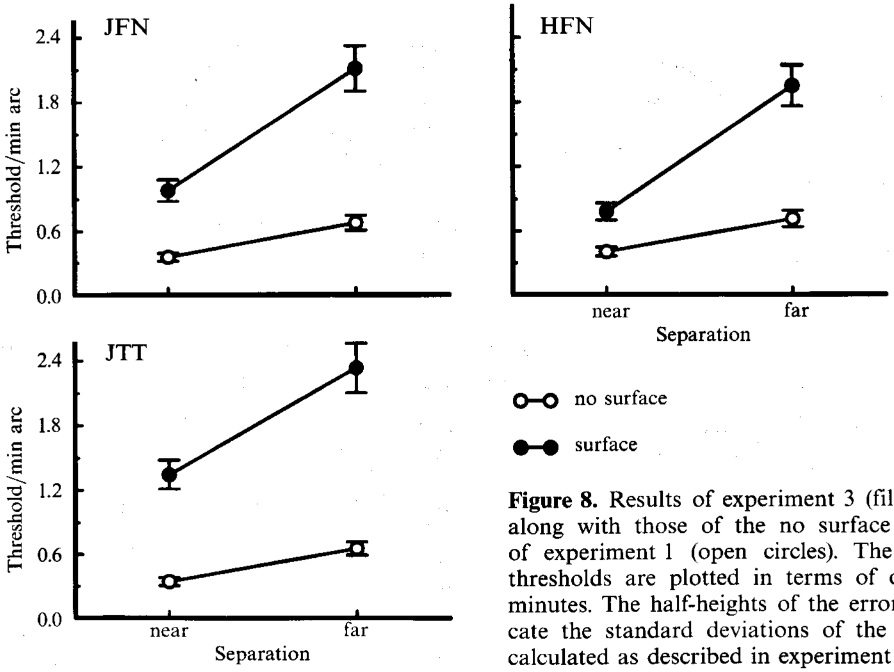

**Figure 8.** Results of experiment 3 (filled circles), along with those of the no surface conditions of experiment 1 (open circles). The observers' thresholds are plotted in terms of disparity in minutes. The half-heights of the error bars indicate the standard deviations of the thresholds, calculated as described in experiment 1.

**Table 3.** Results of experiment 3. Weber fractions expressed as a percentage of the standard.

| | Near | Far |
|---|---|---|
| JFN | 8.7 | 18.7 |
| HFN | 6.4 | 15.9 |
| JTT | 10.8 | 18.7 |

limited to local neighborhoods. This finding agrees with those of Todd and Reichel (1989) and Koenderink and van Doorn (1995), who found that observers' order judgments on a surface defined by shading and texture (no stereo) were greatly affected by the magnitude of image separation.

Perhaps the most obvious effect in figure 8 is the remarkable deterioration that occurred when observers made judgments about the ordering in depth of local regions on smooth surfaces. For any given amount of image separation, the observers needed a much larger disparity to reliably discriminate depth order when surfaces were present. It is important to keep in mind that the presence of an intervening surface does not alter the relative disparities of any given pair of points in space. This result is similar to the one described by Brookes and Stevens (1989). They found that the ability to perceive the ordering of two points in depth was affected by the presence of stereoscopic dotted surfaces. However, in their displays, the two points were placed on two different planar surfaces that were not connected to each other (ie there were large discontinuities in depth between adjacent planes). We find in the current study that observers do not exhibit a precise knowledge of depth order for separated regions on the same object, even when the 3-D structure is multiply defined by stereoscopic disparity, image shading, and texture density.

## 5 Discussion

A common belief among many researchers in visual perception is that the relative depths of points in the environment are a fundamental component of our perceptual representations of 3-D shape, and that the primary source of information about relative depth is provided by binocular disparity. Most existing theoretical models for the computational analysis of 3-D structure from disparity are designed to produce a particular form of data structure called a depth map, that encodes the distance between each visible surface region and the point of observation. Thus, in order to investigate the psychological validity of this approach, the research described in the present article was designed to measure the precision and accuracy of perceived depth relationships under a variety of conditions.

One important issue addressed in these studies involves a distinction between judgments of depth order and judgments of depth intervals. It has been demonstrated in previous research that observers can exhibit hyperacuity at detecting the relative disparity of targets presented near the fixation plane (eg Ogle 1953; Westheimer and McKee 1978), but such judgments do not require an explicit representation of depth or disparity magnitudes. When observers are asked to discriminate depth intervals, as opposed to depth order, the precision of their judgments is significantly reduced. This can be seen most clearly by comparing the results of experiments 1 and 2. For judgments of depth order in experiment 1, the average Weber fraction for the three observers was 3.1% for near separations, but the thresholds for interval judgments were as high as 10.8% in experiment 2.

There are other important aspects of scene context that can also influence the precision of observers' depth discriminations. Research on binocular stereopsis with isolated stimulus elements (lines, bars, spots) in the visual field has sometimes been dismissed as ecologically invalid because visible targets in the natural environment are generally attached to continuous surfaces that provide redundant sources of information such as shading, texture, or motion (eg see Gibson 1950). In one respect, the results of the present experiments could be interpreted to support this argument. There are large differences in the precision of ordinal depth judgments between targets that appear to be floating in space and those that are presented on smoothly curved surfaces defined by stereoscopic disparity, shading, and texture. However, the direction of this difference is exactly the opposite of what one would expect from an ecological perspective.

Positioning targets on a smoothly curved surface can dramatically impair performance, such that observers' discrimination thresholds are two to four times larger than what would otherwise be the case for targets that are presented in isolation. This effect of surfaces appears to be a general one that occurs when one is asked to make any type of judgment about the depths of separated locations in space. It occurs not only for the ordinal judgments used in experiment 3, but also exists when observers perform discriminations of depth intervals (Norman and Todd 1996). Figure 9 shows both the results of the current experiments and those of Norman and Todd (1996), who used a standard depth interval magnitude (2.7 cm) similar to that used in the present set of experiments (3.0 cm). Observers exhibit the greatest precision when they are asked to make depth order judgments for isolated points floating in space. In contrast, judgments about the magnitudes of depth intervals on smooth surfaces are much less reliable.

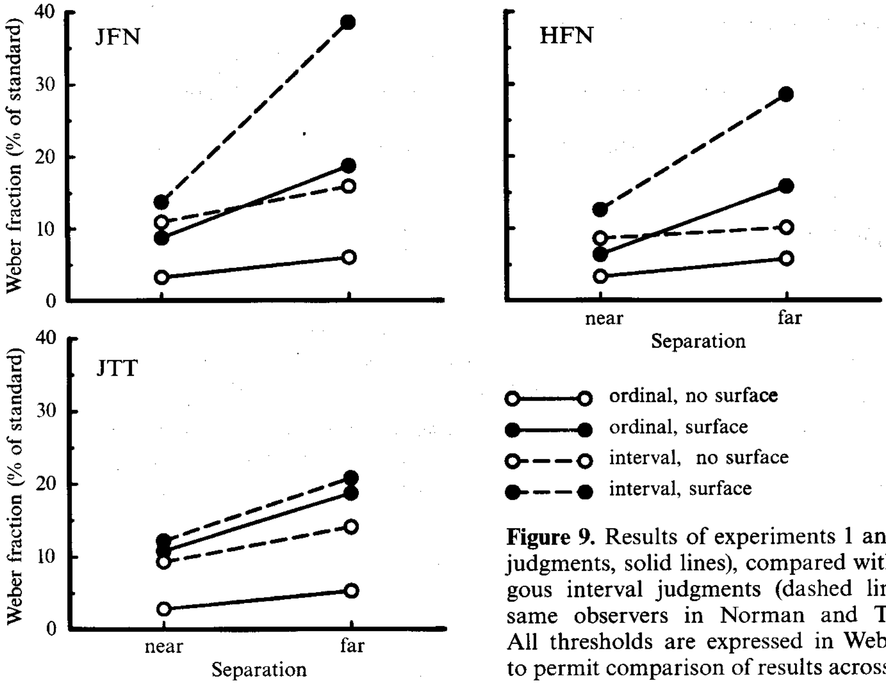

**Figure 9.** Results of experiments 1 and 3 (ordinal judgments, solid lines), compared with the analogous interval judgments (dashed lines) for the same observers in Norman and Todd (1996). All thresholds are expressed in Weber fractions to permit comparison of results across conditions.

One possible explanation why our knowledge of depth is more precise for targets in space and less precise on surfaces is that the perceptual representations of smoothly curved surfaces are primarily concerned with higher-order differential properties such as orientation or curvature. This hypothesis was first suggested by Rogers and Cagenello (1989), who showed that the second derivatives of the disparity field remain invariant with viewing distance, whereas zero-order disparities do not. Thus, by representing curvature rather than depth, the visual system could potentially avoid the stereo scaling problem. There are, however, some important consequences of this approach. If surfaces were perceptually represented as orientation or curvature maps, then in order to compute the relative depth between two designated points it would be necessary to integrate along a path that connects the two separated surface regions. Because errors in this process would tend to accumulate over the length of the path, the precision of observers' depth judgments would be expected to diminish with increased spatial separation. The results of the present experiments and the earlier findings of Norman and Todd (1996) provide clear empirical evidence to confirm this prediction.

Although the discussion thus far has focused on the ability of observers to discriminate interval and ordinal depth relationships, it is also interesting to consider in this context the results obtained with the use of magnitude estimation or matching paradigms. There has been a large body of research reported in the literature over the past century to show that the perceived relief of an object in depth can be systematically distorted, and that it does not remain invariant over changes in viewing distance, even for real objects that are viewed directly in a natural environment (eg Heine 1900; Thouless 1931; Baird and Biersdorf 1967; Wagner 1985; Loomis et al 1992; Norman et al 1996). Taken together, the available empirical evidence suggests that the relative depths among points in a scene may be much less important to our perceptual representations of 3-D form than is generally assumed.

**Acknowledgements.** This research was supported by Air Force Office of Scientific Research grant #F49620-93-1-0116 to James Todd. We are most grateful to four anonymous reviewers for their insightful comments on an earlier draft of the article.

**References**

Baird J C, Biersdorf W R, 1967 "Quantitative functions for size and distance judgments" *Perception & Psychophysics* **2** 161 – 166

Blakemore C, 1970 "The range and scope of binocular depth discrimination in man" *Journal of Physiology (London)* **211** 599 – 622

Brookes A, Stevens K A, 1989 "Binocular depth from surfaces versus volumes" *Journal of Experimental Psychology: Human Perception and Performance* **15** 479 – 484

Cormack R, Fox R, 1985 "The computation of disparity and depth in stereograms" *Perception & Psychophysics* **38** 375 – 380

Enright J T, 1991 "Stereo-thresholds: simultaneity, target proximity and eye movements" *Vision Research* **31** 2093 – 2100

Enright J T, 1996 "Sequential stereopsis: a simple demonstration" *Vision Research* **36** 307 – 312

Fernberger S W, 1931 "On absolute and relative judgments in lifted weight experiments" *American Journal of Psychology* **43** 560 – 578

Foley J M, 1978 "Primary distance perception", in *Handbook of Sensory Physiology* volume VIII *Perception* Eds R Held, H W Leibowitz, H L Teuber (New York: Springer) pp 181 – 213

Foster D H, Bischof W F, 1991 "Thresholds from psychometric functions: superiority of bootstrap to incremental and probit variance estimators" *Psychological Bulletin* **109** 152 – 159

Gibson J J, 1950 *The Perception of the Visual World* (Boston, MA: Houghton-Mifflin)

Heine L, 1900 "Über Orthoskopie oder über die Abhängigkeit relativer Entfernungsschätzungen von der Vorstellung absoluter Entfernung" *Albrecht von Græfe's Archiv für Ophthalmologie* **51** 563 – 572

Helmholtz H von, 1867/1962 *Treatise on Physiological Optics* volume 3 (New York: Dover, 1962); English translation by J P C Southall for the Optical Society of America (1925) from the 3rd German edition of *Handbuch der physiologischen Optik* (Hamburg: Voss, 1909; 1st edition Leipzig: Voss, 1867)

Koenderink J J, Doorn A J van, 1995 "Relief: pictorial and otherwise" *Image and Vision Computing* **13** 321 – 334

Longuet-Higgins H C, 1982 "The role of the vertical dimension in stereoscopic vision" *Perception* **11** 377 – 386

Loomis J M, Da Silva J A, Fujita N, Fukusima S S, 1992 "Visual space perception and visually directed action" *Journal of Experimental Psychology: Human Perception and Performance* **18** 906 – 921

McKee S P, Levi D M, Bowne S F, 1990 "The imprecision of stereopsis" *Vision Research* **30** 1763 – 1779

Norman J F, Todd J T, 1993 "The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations" *Perception & Psychophysics* **53** 279 – 291

Norman J F, Todd J T, 1996 "The discriminability of local surface structure" *Perception* **25** 381 – 398

Norman J F, Todd J T, Perotti V J, Tittle J S, 1996 "The visual perception of three-dimensional length" *Journal of Experimental Psychology: Human Perception and Performance* **22** 173 – 186

Norman J F, Todd J T, Phillips F, 1995 "The perception of surface orientation from multiple sources of optical information" *Perception & Psychophysics* **57** 629 – 636

Ogle K N, 1950 *Researches in Binocular Vision* (Philadelphia, PA: W B Saunders)

Ogle K N, 1953 "Precision and validity of stereoscopic depth perception from double images" *Journal of the Optical Society of America* **43** 906 – 913

Ogle K N, 1956 "Stereoscopic acuity and the role of convergence" *Journal of the Optical Society of America* **46** 269 – 273

Ogle K N, 1958 "Present status of our knowledge of stereoscopic vision" *AMA Archives of Ophthalmology* **60** 755 – 774

Rogers B J, Cagenello R B, 1989 "Disparity curvature and the perception of three-dimensional surfaces" *Nature (London)* **339** 135 – 137

Thouless R H, 1931 "Phenomenal regression to the real object. 1" *British Journal of Psychology* **21** 339 – 359

Thurstone L L, 1927 "A law of comparative judgment" *Psychological Review* **34** 273 – 286

Todd J T, Mingolla E, 1983 "Perception of surface curvature and direction of illumination from patterns of shading" *Journal of Experimental Psychology: Human Perception and Performance* **9** 583 – 595

Todd J T, Mingolla E, 1984 "The simulation of curved surfaces from patterns of optical texture" *Journal of Experimental Psychology: Human Perception and Performance* **10** 734 – 739

Todd J T, Norman J F, 1991 "The visual perception of smoothly curved surfaces from minimal apparent motion sequences" *Perception & Psychophysics* **50** 509 – 523

Todd J T, Norman J F, 1995 "The visual discrimination of relative surface orientation" *Perception* **24** 855 – 866

Todd J T, Reichel F D, 1989 "Ordinal structure in the visual perception and cognition of smoothly curved surfaces" *Psychological Review* **96** 643 – 657

Wagner M, 1985 "The metric of visual space" *Perception & Psychophysics* **38** 483 – 495

Westheimer G, 1979 "Cooperative neural processes involved in stereoscopic acuity" *Experimental Brain Research* **36** 585 – 597

Westheimer G, McKee S P, 1978 "Stereoscopic acuity for moving retinal images" *Journal of the Optical Society of America* **68** 450 – 455

Westheimer G, McKee S P, 1979 "What prior uniocular processing is necessary for stereopsis?" *Investigative Ophthalmology & Visual Science* **18** 614 – 621

Westheimer G, McKee S P, 1980 "Stereogram design for testing local stereopsis" *Investigative Ophthalmology & Visual Science* **19** 802 – 809

Wever E G, Zener K E, 1928 "The method of absolute judgment in psychophysics" *Psychological Review* **35** 466 – 493

Wheatstone C, 1838 "Contributions to the physiology of vision—Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision" *Philosophical Transactions of the Royal Society of London* **128** 371 – 394

Wright W D, 1951 "The role of convergence in stereoscopic vision" *Proceedings of the Physical Society* **64** 289 – 297